

Open Research Online

The Open University's repository of research publications and other research outputs

Inverting Khintchine's relationship and generating length biased data

Journal Item

How to cite:

Jones, M. C. (2019). Inverting Khintchine's relationship and generating length biased data. *Statistics & Probability Letters*, 154, article no. 108539.

For guidance on citations see [FAQs](#).

© 2019 Elsevier



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Accepted Manuscript

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.1016/j.spl.2019.06.015>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

Accepted Manuscript

Inverting Khintchine's relationship and generating length biased data

M.C. Jones

PII: S0167-7152(19)30176-2
DOI: <https://doi.org/10.1016/j.spl.2019.06.015>
Reference: STAPRO 8539

To appear in: *Statistics and Probability Letters*

Received date: 27 March 2019
Revised date: 17 June 2019
Accepted date: 18 June 2019

Please cite this article as: M.C. Jones, Inverting Khintchine's relationship and generating length biased data. *Statistics and Probability Letters* (2019), <https://doi.org/10.1016/j.spl.2019.06.015>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Inverting Khintchine's Relationship and Generating Length Biased Data

M.C. Jones

School of Mathematics and Statistics, The Open University, Walton Hall, Milton Keynes MK7 6AA, U.K.

ABSTRACT

If $X > 0$ follows a distribution with decreasing density, then Khintchine's theorem states that it has the same distribution as $U \times S$ where U and $S > 0$ are independent, U following the uniform distribution on $(0, 1)$. In this letter, an explicit function of X and independent $V \sim U(0, 1)$ is discovered which has the same distribution as S . This result is then used to find an explicit function of two independent uniform random variables which follows the length biased form of a general distribution on \mathbb{R}^+ with finite mean.

Keywords:

Decreasing density

Gibbs sampling

Khintchine's theorem

Uniform random variables

1. Introduction

This letter concerns two previously unresolved questions that turn out to have answers emanating from the same source. The two questions are as follows:

- (1) If $X > 0$ follows a distribution D with decreasing density d , say, then it has the same distribution as $U \times S$ where U and $S > 0$ are independent, U following the uniform distribution on $(0, 1)$, written $U \sim U(0, 1)$. This is a version of Khintchine's theorem (Khintchine, 1937, Feller, 1971, Jones, 2002). Is there a way of “inverting” the distributional relationship associated with Khintchine's theorem so that some function of $X \sim D$ and independent $V \sim U(0, 1)$ has the same distribution as S ? The answer is certainly not to set $S = X/V$, of course.

- (2) Let f be the density of a distribution F on \mathbb{R}^+ with finite mean μ . If $Y > 0$ follows the corresponding length-biased distribution with density $g(y) = yf(y)/\mu$ (see Brown, 2006, for much on length-biased distributions), is there an explicit function of independent uniform random variables by which realisations of Y can be generated? (Here, an alternative answer is sought to the probability integral transformation $Y = G^{-1}(U)$ where it is necessary to invert $G(y) = \{yF(y) - W(y)\}/\mu$ where $W(y) = \int_0^y F(w)dw$ is the iterated distribution function associated with F , e.g. Bassan et al., 1999, this might be difficult.)

The source of positive answers to both of these questions is iteration around conditional distributions as in Gibbs sampling. Gibbs sampling is therefore discussed in the relevant, bivariate, context in Section 2 before being used to give answers to Questions (1) and (2) in Sections 3 and 4, respectively. The result of Section 4 is, in turn, an application of the result of Section 3. Some further remarks complete the letter in Section 5.

2. Gibbs sampling for univariate decreasing distributions

Let X follow a distribution D with density d on support \mathcal{S} . Then, as is well known, $X \sim D$ is the x -marginal distribution of the bivariate distribution of $\{X, Y\}$ which has density

$$f_{X,Y}(x, y) = \mathbb{I}_{\mathcal{S}}(x) \mathbb{I}_{(0, d(x))}(y);$$

here, $\mathbb{I}_A(z)$ is the indicator function that $z \in A$ i.e. $\{X, Y\}$ is uniformly distributed over the region between \mathcal{S} and the horizontal axis. The conditional distribution of $Y|X = x$ is, immediately, the uniform distribution on $(0, d(x))$:

$$f_{Y|X=x}(y) = \frac{1}{d(x)} \mathbb{I}_{(0, d(x))}(y).$$

Now confine attention to $\mathcal{S} = (0, b)$ where $b > 0$ and will often be ∞ , and take d to be continuous and decreasing for all $0 < y < b$ ($d(b)$ will usually be 0). It follows that d has an inverse, $d^{-1}(y)$, for $d(b) \leq y < d(0)$, and hence that the conditional distribution of $X|Y = y$ is the uniform distribution on $(0, \mathcal{D}(y))$,

$$f_{X|Y=y}(x) = \frac{1}{\mathcal{D}(y)} \mathbb{I}_{(0, \mathcal{D}(y))}(x),$$

where

$$\mathcal{D}(y) = \begin{cases} d^{-1}(y) & \text{if } y \geq d(b), \\ b & \text{if } y < d(b). \end{cases} \quad (1)$$

Note that if $b = \infty$ and/or $d(b) = 0$, $\mathcal{D} = d^{-1}$.

Since both “full” conditional distributions of $f_{X,Y}$ are therefore readily available, indeed both are uniform distributions, realisations of random variables from this distribution could be obtained through Gibbs sampling. In the following algorithm, each $U_i \sim U(0, 1)$ independently of other U_i ’s.

1. Choose a value of $0 < x_0 < b$ and set $k = 1$.
2. Take Y_k to be uniformly distributed on $(0, d(x_{k-1}))$; that is, set $Y_k = U_{2k-1} \times d(x_{k-1})$, with realisation y_k .
3. Take X_k to be uniformly distributed on $(0, \mathcal{D}(y_k))$; that is, set $X_k = U_{2k} \times \mathcal{D}(y_k)$, with realisation x_k .
4. Change k to $k + 1$ and iterate steps 2 and 3 until convergence.

This is a valid, if perhaps rather inefficient and often unnecessary, way of generating random variables following the distribution with decreasing density d on support $(0, b)$: discard all y values and, after a suitable burn-in period, retain a suitably thinned series of x values. The algorithm may be of interest in its own right (for one relevant result, see Section 5) but here it is used to form the basis of answers to Questions (1) and (2) posed in Section 1.

3. Inverting Khintchine’s relationship

If x_0 in the Gibbs sampling algorithm of Section 2 were actually a realisation of $X \sim D$, then so would be X_1 (and X_2, X_3, \dots). That is, X_1 could be written in the form

$$X_1 = U_2 \mathcal{D}\{U_1 d(X)\}. \quad (2)$$

But this is precisely of the form $X_1 = U \times S$ where $U = U_2 \sim U(0, 1)$ is independent of

$$S = \mathcal{D}\{V d(X)\}. \quad (3)$$

Formula (3) for S is, as requested in Question (1) of Section 1, a function of $X \sim D$ and independent $V = U_1 \sim U(0, 1)$. This is summarised in Result 3.1.

RESULT 3.1

Let $X > 0$ follow a distribution D with continuous and decreasing density d for all $0 < x < b$, where b may be ∞ . Then, the following relationships hold:

- $X = US$ where $U \sim U(0, 1)$ independently of $S \sim C$; and
- $S = \mathcal{D}\{V d(X)\}$ where $V \sim U(0, 1)$ independently of $X \sim D$.

The relationship between C and D can be derived as follows

$$\begin{aligned} P(S \leq s | X = x) &= P(U < d(b)/d(x)) \mathbb{I}_{\{b\}}(s) + P(U \geq d'(s)/d'(x)) \mathbb{I}_{(0,b)}(s) \\ &= \frac{d(b)}{d(x)} \mathbb{I}_{\{b\}}(s) + \left\{ 1 - \frac{d(s)}{d(x)} \right\} \mathbb{I}_{(0,b)}(s) \end{aligned}$$

so that

$$\begin{aligned} C(s) = P(S \leq s) &= \int_0^b d(b) dx \mathbb{I}_{\{b\}}(s) + \int_0^s \{d'(x) - d(s)\} dx \mathbb{I}_{(0,b)}(s) \\ &= b d(b) \mathbb{I}_{\{b\}}(s) + \{D(b) - s d'(s)\} \mathbb{I}_{(0,b)}(s). \end{aligned}$$

When $b d(b) = 0$ and d is differentiable, $c(s) = C'(s) = -s d'(s)$, which is a familiar relationship forming part of Khintchine's theorem as often stated.

EXAMPLE 3.1

- (a) Let $d(x) = \lambda e^{-\lambda x}$ be the density of the exponential distribution with parameter $\lambda > 0$, $M(\lambda)$ say. In this case, $c(x) = \lambda^2 x e^{-\lambda x}$ is the gamma distribution with parameters 2 and λ , written $G(2, \lambda)$. Also, $\mathcal{D}(y) = d^{-1}(y) = -(1/\lambda) \log(y/\lambda)$ so that, using (3), $S = X - (1/\lambda) \log U$. That $S \sim C$ is true is because X and $-(1/\lambda) \log U$ are independent $M(\lambda)$ random variables, and S is their convolution.
- (b) Let $d(x) = \gamma/(1+x)^{\gamma+1}$, $\gamma > 0$, be the density of the unit-scale Lomax (or $(1/\gamma) \times F_{2,2\gamma}$) distribution. In this case, $c(x) = \gamma(\gamma+1)x/(1+x)^{\gamma+2}$ is the density of the $F_{4,2\gamma}$ distribution, scaled by $2/\gamma$. Also, $\mathcal{D}(y) = d^{-1}(y) = (\gamma/y)^{1/(\gamma+1)} - 1$ so that, using (3) again,

$$S = U^{-1/(\gamma+1)}(1 + X) - 1.$$

This gives an interesting new relationship between the $F_{2,2\gamma}$ and $F_{4,2\gamma}$ (or $F_{2,\nu}$ and $F_{4,\nu}$) distributions.

- (c) As an example with “ $b d(b) \neq 0$ ”, let $d(x) = \alpha x^{\alpha-1} \mathbb{I}_{(0,1)}(x)$, for $0 < \alpha < 1$. Now, $C(s) = \alpha \mathbb{I}_{\{1\}}(s) + (1-\alpha)s^\alpha \mathbb{I}_{(0,1)}(s)$, that is, $S = 1$ with probability α and is distributed as D with probability $1-\alpha$. In this case,

$$S = \mathcal{D}(Vd(X)) = \begin{cases} V^{-1/(1-\alpha)}X & \text{if } V \geq X^{1-\alpha}, \\ 1 & \text{if } V < X^{1-\alpha}. \end{cases}$$

4. Generating length biased data

Consider distributions whose decreasing, differentiable, densities d are defined to be proportional to decreasing, differentiable, proper survival functions on $(0, b)$ where b may be ∞ ; such a survival function, \bar{F} say, with density f , must have finite mean μ , and then $d(x) = \bar{F}(x)/\mu$. Note that $d(0) = 1/\mu < \infty$. Such a d is a special case of the more general d 's considered above. The corresponding density $c(y) = -yd'(y) = yf(y)/\mu = g(y)$ is nothing other than the length-based density associated with f . Setting $d(x) = \bar{F}(x)/\mu$ in (3) therefore yields, for $X \sim \bar{F}/\mu$ and independent $V \sim U(0, 1)$,

$$Y = \bar{F}^{-1}\{V\bar{F}(X)\} \sim G \quad (4)$$

where G is the distribution function associated with density g . Result 4.1 below follows from (4) by the probability integral transformation, noting that \bar{W}/μ is the survival function associated with density \bar{F}/μ , where the iterated survival function $\bar{W}(x)$ is defined by $\int_x^b \bar{F}(w)dw$ (e.g. Bassan et al., 1999). Note that f in Result 4.1 is *not* constrained to be decreasing.

RESULT 4.1

Let f be a density on $(0, b)$ which has finite mean μ even when $b = \infty$. If $U \sim U(0, 1)$ and, independently, $V \sim U(0, 1)$, then

$$Y = \bar{F}^{-1}[V\bar{F}\{\bar{W}^{-1}(\mu U)\}] \quad (5)$$

follows the length-biased distribution with density $g(y) = yf(y)/\mu$.

That g is the density of Y in (5) can be confirmed by direct manipulations. Most easily,

$$\begin{aligned} \bar{G}(y) &= P(Y \geq y) = P\left(V \leq \bar{F}(y)/\bar{F}\{\bar{W}^{-1}(\mu U)\}\right) \\ &= \bar{F}(y) \int_{\bar{W}(y)/\mu}^1 \frac{1}{\bar{F}\{\bar{W}^{-1}(\mu u)\}} du + \int_0^{\bar{W}(y)/\mu} du \\ &= -\bar{F}(y) \int_{\bar{W}(y)/\mu}^1 (\bar{W}^{-1})'(\mu u) du + \frac{\bar{W}(y)}{\mu} \\ &= -\frac{\bar{F}(y)}{\mu} [\bar{W}^{-1}(\mu) - \bar{W}^{-1}\{\bar{W}(y)\}] + \frac{\bar{W}(y)}{\mu} \\ &= \frac{1}{\mu} \{y\bar{F}(y) + \bar{W}(y)\}. \end{aligned}$$

Differentiation of this formula for $\bar{G}(y)$ gives the required form for the density $g(y)$.

EXAMPLE 4.1

- (a) The exponential distribution is the unique distribution for which \bar{F}/μ and f are the same, and so the manipulations of Example 3.1(a) continue to hold directly: since $\mu = 1/\lambda$, the length-biased distribution is the $G(2, \lambda)$ distribution, $\bar{W}(x) = e^{-\lambda x}/\lambda$ and

$$Y = -(\log U + \log V)/\lambda.$$

This is a standard result relating the $G(2, \lambda)$ distribution and the convolution of exponentials generated by $-(1/\lambda)\log U$ and $-(1/\lambda)\log V$. However, the availability of this formula from the general explicit formula (5) contrasts with an attempt to use the probability integral transformation (involving a single uniform random variable) directly, whose explicitness founders on the need to invert $G(y) = 1 - (1 + \lambda y)e^{-\lambda y}$.

- (b) Consider the unit-scale Lomax distribution when $\gamma > 1$ (so that it has finite mean) as f . In this case, $g(x) = (\gamma - 1)x/(1 + x)^{\gamma+1}$ is the density of the $F_{4,2(\gamma-1)}$ distribution, scaled by $2/(\gamma - 1)$. In this case, $\mu = 1/(\gamma - 1)$, $\bar{F}(x) = 1/(1 + x)^\gamma$, $\bar{F}^{-1}(y) = y^{-1/\gamma} - 1$, $\bar{W}(x) = 1/\{(\gamma - 1)(1 + x)^{\gamma-1}\}$ and $\bar{W}^{-1}(y) = [1/\{(\gamma - 1)y\}^{1/(\gamma-1)}] - 1$. It follows that

$$Y = \frac{1}{U^{1/(\gamma-1)} V^{1/\gamma}} - 1. \quad (6)$$

This gives the scaled $F_{4,2(\gamma-1)}$ random variable Y as a simple function of independent $U \sim U(0, 1)$ and $V \sim U(0, 1)$. Again, however, the relationship — obtained here from a quite different approach — is already known: the standard result that the product of independent $\text{Beta}(a, b)$ and $\text{Beta}(a+b, c)$ random variables is distributed as $\text{Beta}(a, b+c)$ gives the $\text{Beta}(\gamma - 1, 2)$ distribution for the denominator of (6), and the usual relationship between beta and F distributions completes the argument.

5. Further remarks

- There is a version of Gibbs sampling for any continuous distribution on \mathbb{R} . When the density d is unimodal, it also leads to versions of Khintchine's theorem. The simple multiplicative relationship addressed in Section 3 arises also if unimodal d is, in addition, symmetric about zero, in which case $U \sim U(0, 1)$ in Result 3.1 is replaced by $W = 2U - 1 \sim (-1, 1)$.
- The correlation between X and X_1 , given by (2), which is relevant to the later stages of the Gibbs sampling algorithm in Section 2, is considered here when $b = \infty$ for convenience. Provided that $s_2 \equiv \mathbb{E}(X^2) < \infty$, so that $\sigma^2 \equiv \mathbb{V}(X) < \infty$, it is the case that

$$\begin{aligned}
\mathbb{E}(XX_1) &= \mathbb{E}[XUd^{-1}\{Vd(X)\}] = \frac{1}{2}\mathbb{E}[Xd^{-1}\{Ud(X)\}] \\
&= \frac{1}{2} \int_0^1 \int_0^\infty xd^{-1}\{vd(x)\} d(x) dx dv \\
&= \frac{1}{2} \int_0^1 \int_{d^{-1}\{vd(0)\}}^\infty (d^{-1})' \left\{ \frac{d(w)}{v} \right\} w \frac{d(w)}{v} d^{-1} \left\{ \frac{d(w)}{v} \right\} \frac{d'(w)}{v} dw dv \\
&= \frac{1}{2} \int_0^\infty wd'(w) \int_{d(w)}^{d(0)} (d^{-1})'(c) d^{-1}(c) dc dw \\
&= -\frac{1}{4} \int_0^\infty w^3 d'(w) dw = \frac{3}{4} \int_0^\infty w^2 d(v) dw = \frac{3s_2}{4}.
\end{aligned}$$

It follows that $\text{Cov}(X, X_1) = (3s_2/4) - \mu^2 = \sigma^2 - (s_2/4)$ so that

$$\text{Corr}(X, X_1) = 1 - \frac{R}{4} \quad \text{where} \quad R = \frac{s_2}{\sigma^2}.$$

Now, clearly, $R > 1$ and, for a decreasing distribution, Johnson and Rogers (1951) showed that $\sigma^2 > \mu^2/3$, it follows that $\sigma^2 > s_2/4$ and hence $R < 4$. So, finally, $0 < \text{Corr}(X, X_1) < 3/4$.

References

- Bassan, B., Denuit, M. and Scarsini, M. (1999) Variability orders and mean differences. *Statistics and Probability Letters*, 45, 121–130.
- Brown, M. (2006) Exploiting the waiting time paradox. *Probability in the Engineering and Informational Sciences*, 20, 195–230.
- Feller, W. (1971) *An Introduction to Probability Theory and Its Applications*, (Vol. 2). New York: Wiley.
- Johnson, N.L. and Rogers, C.A. (1951) The moment problem for unimodal densities. *Annals of Mathematical Statistics*, 22, 433–439.
- Jones, M.C. (2002) On Khintchine's theorem and its place in random variate generation. *The American Statistician*, 56, 304–307.
- Khintchine, A.Y. (1938) On unimodal distributions. *Izvestiya Nauchno-Issledovatel'skogo Instituta Matematiki i Mekhaniki*, 2, 1–7.